THE EMBO JOURNAL

Manuscript EMBO-2011-79465

Multiple factors dictate target selection by Hfq-binding small RNAs

Chase L. Beisel, Taylor B. Updegrove, Ben J. Janson and Gisela Storz

Corresponding authors: Chase L. Beisel and Gisela Storz, Eunice Kennedy Shriver National Institute of Child Health and Human Development

Review timeline:

Submission date:	12 September 2011
Editorial Decision:	21 October 2011
Revision received:	26 January 2012
Editorial Decision:	06 February 2012
Revision received:	07 February 2012
Accepted:	08 February 2012

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 21 October 2011

Thank you for submitting your manuscript for consideration by The EMBO Journal. I have now received the final report from the three referees who evaluated your study and I enclose their comments below. As you will see they are in general positive regarding the analysis of parameters that define sRNA targets but require some further experimental analysis. After careful consideration of the additional requested work, they seem to be important to strengthen the current data and main conclusions in the manuscript and therefore should be addressed. Given the support form the referees I would like to invite you to submit a revised version of the manuscript.

I should remind you that it is EMBO Journal policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript. When you submit a revised version to the EMBO Journal, please make sure you upload a letter of response to the referees' comments. Please note that when preparing your letter of response to the referees' comments that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process initiative, please visit our website: http://www.nature.com/emboj/about/process.html

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1

Beisel et al. have addressed intriguing questions regarding the base-pairing and target site structure requirements of Hfq-associated small regulatory sRNAs (sRNAs), using Spot42 sRNA as a model. The starting point of their analysis was a biocomputational search for mRNA targets of Spot42 in E. coli, in which the authors discovered that querying only the unstructured regions of the molecule (instead of the whole sRNA sequence) significantly improved the accuracy of base-pairing predictions. Next, the authors went on to investigate false-positive targets, i.e. those predicted to form sufficiently long interaction but which failed validation of regulation by Spot42 in vivo. This discovered three main criteria for Spot42-mediated base-pairing: 1) accessibility of the targeting regions 2) stability of the duplex and 3) binding and positioning of the RNA chaperone Hfq on the mRNA.

This is a very timely and important manuscript and the first work to go step-by-step and try to understand why some mRNAs are regulated by a given sRNA, whereas others are not despite as convincing-looking in silico predictions of base-pairing. The manuscript is very well written and the rationale of the experiments is clearly defined. The criticism below should help the authors improve their manuscript prior to publication in EMBO J.

Major criticism:

1. Figure 2: The results for Spot42-mediated repression of gltA::lacZ and srlA::lacZ (boxes on the right hand side in panels A and B, respectively) look pretty much the same as in Beisel et al. 2011 Mol Cell paper. I understand that these are important reference values, so they do not compromise the novelty of the entire figure. If it is the same data indeed, this should be indicated in the figure legend.

2. Figure 3A: There is really not much support for the interactions of Spot42 with sthA or ascF, not to mention the predicted circularization of these mRNAs (as originally proposed by the Belfort lab for DsrA sRNA and several of its targets). The authors should provide evidence that such mRNA circularization is essential for regulation, or otherwise omit these targets.

3. Figure 3B-C: Previous footprint assays demonstrated that Spot42 binds galK mRNA at three distinct sites (Møller et al. 2002 Genes & Dev), which adds confidence that Spot42 makes double contact with nanC as well. However, given the large excess of unlabelled RNA used in these assays (50- and 500-fold), it is impossible to say whether the observed protection occurs in a binary complex-one Spot42 molecule bound to one nanC molecule-or a ternary complex in which Spot42 molecules independently of each other protect the two different sites in nanC. This should be tested by gel retardation experiments, and probing should be performed with concentrations closer to the Kd of the Spot42-nanC interaction.

4. According to Figure 3A, nanC is binding Spot42 at positions 1-11 and 49-56. However, the T1 protection assay does not show protection at a 50-fold excess of nanC and when a 500-fold excess is supplied, protection is observed in parts that should not interact with nanC (e.g. G14, G18, G21, G61, G62 and G65). Furthermore, the authors have previously shown that mutation in Spot42 region III does not impair repression of a nanC::lacZ reporter, whereas mutation of Spot42 region I had a significant effect (Beisel et al. 2011 Mol Cell). This needs to be clarified.

5. How does Hfq impact the proposed Spot42 interaction with multiple regions in targets?

6. Part of the data in Figure 4A is unclear. The parental srlA::lacZ fusion is regulated 2.8-fold in the presence of Spot42 (as shown in Fig. 2B). Base-pairing occurs via region II of Spot42, and mutation of this region impairs regulation (Beisel et al. 2011 Mol Cell). Now the authors extend the base-pairing with the 5'UTR of srlA (srlA+III::lacZ) and observe strongly increased regulation (27-fold

repression), which is attributed to new base-pairing with region III of Spot42. However, subsequent mutation of Spot42 region III reverted repression not to 2.8-fold but to 6.3-fold. Moreover, the srlA-II+III::lacZ reporter (mutation of srlA opposite to part II of Spot42) was still down-regulated by mutant pSpot42-III. In addition, in their previous article the authors showed that mutation of Spot42 region III also impacted on the repression of the parental srlA::lacZ reporter (Beisel et al. 2011 Mol Cell). I know that none of these experiments and mutations is expected to give all-or-none results but this really needs to be explained much better.

7. Figures 1, 2, 4, 5 and 6: the mutations introduced in the target mRNA fusions often impact on the basal β-galactosidase levels of the reporters. This indicates that translation activity and/or RNA stability greatly differs from the parental (wild-type) fusions. Since the authors use "fold-regulation" as a read-out for the base-pairing activity of Spot42, they should show or argue why they ignore the possibility that altered translation rates (or ribosome occupancy) are responsible for the observed differences in regulation. Take the srlA-entB::lacZ reporter in Figure 6G, which has a basal β-galactosidase activity of ~8,000 MU and is regulated 4.2-fold by Spot42. By contrast, the entB+I::lacZ reporter has a basal activity of only ~530 MU (more than a log difference) and shows only 1.7-fold regulation, despite longer base-pairing.

8. Page 11, first para, and Figure 5: Does fusion of the 5' end of srlA to the predicted targets also impart association with Hfq in vivo? The in vitro experiments are suggestive, but looking into enrichment of the "repaired" target fusion in coIP with Hfq in vivo/lysates, for example, using qRT-PCR, would be more convincing. This is a simple and straight-forward experiment. With regards to the in vitro work in Figure 5A, does a trimeric srlA-usg/Hfq/Spot42 when the sRNA is added as well?

Minor points:

9. Regarding guiding principles for base-pairing sRNAs, one would also mention a paper by Richter et al. 2010 Bioinformatics 26(1):1-5 in which a similar strategy discovered two novel targets of Yfr1 sRNA.

10. Page 6, end of first para: ideally, one would like to see changes in abundance of the respective native target protein before concluding that some targets are false-positive. It could be that regulation is at the level or translation only, and/or the used reporters are just not suitable. Briefly mention these possibilities to caution.

11. Page 8, first para: again, regulation can be more than mRNA changes, which is what the cited papers assessed mainly.

12. Figure 1A: the structure model of Spot42 needs labels to interpret the proposed base-pairing interactions.

13. Figure 1: puuE is the third gene in an operon. Does Spot42 regulate the upstream puuC and pub genes as well?

14. Table II: the authors argue on page 7, 2nd para, that 10/15 predicted targets were repressed. However, Table II includes only nine (nanT, nanC, paaK, atoD, ascF, galK, srlA, fucP and glpF).

15. Figure 2B: for the srlA_L::lacZ extended base-pairing into the structured Spot42 sequence improved base-pairing from 2.8-fold to 5.5-fold while srlA_R::lacZ (extending into the unstructured regions of Spot42) improved base-pairing to 12.5- fold. Is this really due to different stabilities of complexes or to different structural constraints (the Spot42-srlA_L interaction will require three additional bulged residues).

16. Fig. 6G: the increase in β-galactosidase activity of the entB+I::lacZ reporter when transformed with the Spot42 plasmid (without IPTG) should be explained, specifically because this is crucial for the interpretation of the experiment.

Referee #2

Beisel et al. Multiple factors dictate..

The paper by Beisel et al addresses a practical issue - how to identify targets of bacterial sRNAs based on features that determine an efficient regulatory interaction - as well as the important general question of what these features are. The authors conduct their study in E. coli and use as their model sRNA Spot42. A recent report from the Storz lab had already identified several targets and highlighted their biological relevance. The starting point of this work is a list of Spot 42 targets computationally predicted by a fairly simple algorithm. When analyzed by microarray and reporter fusion techniques, most predictions appeared not to be supported. The authors then proceeded to incorporate several features into the prediction that might be related to regulatory efficiency, such as the need for single-strandedness/ low structure content in the interaction regions of sRNA and target, the length of the base-pairing region, the use of more than one base-pairing region, and a suitably spaced Hfq binding site in the targeted mRNA. It is also argued that Hfq binding to the target sequence could interfere with sRNA-promoted inhibition.

Most of the data that test for contributions of each of the features above are derived from well-designed translational fusion (target-lacZ) experiments. Three ss-regions are present in Spot 42. Mutations in these, or in the target, report on base-pairing requirements for inhibition, and compensatory mutations are used to restore lost regulation. Additional mutations were introduced to create longer base-pairing regions, to break interfering structures, or to engineer Hfq binding motifs into target RNAs. Broadly speaking, the paper makes a convincing case for the list of features that "create" a target, or turn a seemingly plausible interaction site into a non-target.

The paper is well-written, the arguments are overall concisely stated and convincingly argued. The conclusions are of broad general interest and fit the scope of EMBO Journal. They address in one body of work what so far were non-systematically scattered inadvertant observations and conclusions in many publications in the sRNA field.

There are a few issues that deserve to be addressed

1) Since the experiments in Figs 5 and 6 involve major mutational changes which could have effects other than the ones desired, and since Hfq binding is here assumed to be a key factor in reactivating silent targets, these experiments should be repeated in an Hfq-deficient strain. This straightforward experiment should also be done for the extended-target fusions in Figure 2, in particular since the Gottesman lab has argued that longer interactions may overcome an Hfq requirement.

2) Two of the three not regulated cases (usg, moeA) are straightforward. The entB case is complicated, and even though the interpretations given may turn out to be correct, there are loose ends. Since competition between Hfq and Spot 42 for the downstream Hfq binding site/target site is assumed, this could easily be tested in vitro. For instance, by analyzing gelshifts of an mRNA fragment lacking the upstream binding site with Spot42 and/or Hfq. And again, an in vivo test in an Hfq mutant strain might help. It might also be worth a few more lines to discuss whether such a model can account for the in vivo effects. For Hfq blocking access to a target site, its occupancy must be very high (most of the mRNAs at any given time bound by Hfq at this site).

3) Fig. 3: Though the reporter gene experiments are convincing, the biochemical analysis of the interaction site is not. Except for the lead cleavages at 2 μM nanC, I find it difficult to see any convincing pattern. In particular, it is unclear why RNAse III was used as a probe for dsRNA regions. V1 has this specificity whereas RNAse III usually cleaves 20 or so base-pair stretches of RNA, far longer than assumed here (Fig 3A). Incidentally, the choice of concentration ranges in this experiment, as stated, comes from the binding curves in Suppl. F.3. There, the mid-point (50% labeled RNA bound) differs in the reciprocal experiments by maybe 4-fold. This is odd, to say the least, unless a binding mode different from 1:1 is assumed.

4) For many of the fusion constructs, adding for example the srlA leader region changes lacZ activities substantially (e.g. moeA down 5x, entB up 8x). Is this an indication of changed mRNA levels, or translational activity?

Referee #3

In this study, Beisel et al. addressed a question what are responsible for target selection among mRNAs predicted to base pair with Hfq-binding sRNAs using Spot42 of E. coli as a model. The authors started to identify potential targets of Spot42 using algorithsms. Then, they performed a series of mutational analysis of predicted targets by generating lacZ translational fusion and by lacZ assay to assess Spot42 regulation. Major results/observations are: 1) base-pairing prediction using only unstructured regions of Spot42 is quite effective to identify the targets; 2) mutational analysis supports that the predicted base pairings are indeed important for the regulation by Spot42; 3) increasing the extent of base pairing in unstructured region of weakly regulated targets strengthen the regulation; 4) some of non-targets predicted to base pair lack an Hfq-binding site and the introduction of an Hfq-binding site converted the non-target to target; 5) the putative base pairing regions of other non-targets are folded or contain an Hfq-binding site that potentially occlude the targeting site and these non-targets became regulated by Spot42 by disrupting the secondary structure or the Hfq-binding site. Based on these results, the authors conclude that multiple factors are involved in target selection.

The experiments are well designed, and executed thoroughly and carefully. The data and arguments are mostly clear and convincing to support the conclusion. The work certainly contributes to understanding the action of bacterial base-pairing sRNAs in particular the principles underlying the target selection of Hfq-binding sRNAs. I only have a few comments/questions.

1) Many targets predicted by "TargetRNA" using the full-length Spot42 seem to base pair with the unstructured regions of Spot42 (Table 1). Are these unstructured regions different from the regions I, II, and III or do they overlap each other?
2) The mRNA levels of two targets regulated as lacZ fusion by Spot42 are significantly elevated in the Δspf strain lacking Spot42 as expected while those of others are not (Supplementary Fig. 1). What are the possible reasons for this?

1st Revision - authors' response 26 January 2012

We thank the referees for their constructive and in-depth comments, which have helped us improve the manuscript. Responses to each comment are below.

Referee #1

1. Figure 2: The results for Spot 42-mediated repression of *gltA::lacZ* and *srlA::lacZ* (boxes on the right hand side in panels A and B, respectively) look pretty much the same as in Beisel et al. 2011 Mol Cell paper. I understand that these are important reference values, so they do not compromise the novelty of the entire figure. If it is the same data indeed, this should be indicated in the figure legend.

We appreciate the referee's thoroughness in comparing results from this manuscript to our recent report in *Mol Cell*. We affirm that the data in Figure 2 and in our previous work were collected independently.

2. Figure 3A: There is really not much support for the interactions of Spot 42 with *sthA* or *ascF*, not to mention the predicted circularization of these mRNAs (as originally proposed by the Belfort lab for DsrA sRNA and several of its targets). The authors should provide evidence that such mRNA circularization is essential for regulation, or otherwise omit these targets.

We agree that our data for *sthA* and *ascF* do not differentiate between circularization and

other potential mechanisms such as 2:1 base-pairing stoichiometry or 1:1 base-pairing stoichiometry with single-site hybridization (both mechanisms are described on p.15-16 of the Discussion). Since the previous depiction in Figure 3A implies circularization, we have separated the two targeting sites in the figure panel. The new panel does not favor any one mode of interaction over another.

We thank the referee for pointing out the previous prediction by the Belfort lab that DsrA base pairs with some of its mRNA targets in multiple locations. We have added a reference to this original proposal to p. 15:

> **"This parallels the previous prediction that DsrA base pairs near the start codon and the stop codon of a subset of mRNAs (Lease & Belfort, 2000) …"**

3. Figure 3B-C: Previous footprint assays demonstrated that Spot 42 binds *galK* mRNA at three distinct sites (Møller et al. 2002 *Genes & Dev*), which adds confidence that Spot 42 makes double contact with *nanC* as well. However, given the large excess of unlabeled RNA used in these assays (50- and 500-fold), it is impossible to say whether the observed protection occurs in a binary complex – one Spot 42 molecule bound to one *nanC* molecule – or a ternary complex in which Spot 42 molecules independently of each other protect the two different sites in *nanC*. This should be tested by gel retardation experiments, and probing should be performed with concentrations closer to the K_D of the Spot 42-*nanC* interaction.

As suggested by this referee, we repeated the structural probing experiments using unlabeled RNA concentrations closer to the apparent K_D of the Spot42:*nanC* mRNA interaction. Our results showed that both predicted pairing sites were more structured in the presence of unlabeled RNA. While the concentrations of the unlabeled RNA exceeded the concentration of the labeled RNA as pointed out by this referee, this ratio is not important as long as the K_D exceeds the concentration of the labeled RNAs. For these experiments the K_D was ~250 nM and the concentration of the labeled RNAs was ~20 nM.

The stoichiometry of pairing between sRNAs and target mRNAs in vitro and in vivo is an intriguing question that goes beyond the scope of the current work, which focused on whether multiple pairing sites can improve fold regulation.

4. According to Figure 3A, *nanC* is binding Spot 42 at positions 1-11 and 49-56. However, the T1 protection assay does not show protection at a 50-fold excess of *nanC* and when a 500-fold excess is supplied, protection is observed in parts that should not interact with *nanC* (e.g. G14, G18, G21, G61, G62 and G65). Furthermore, the authors have previously shown that mutation in Spot 42 region III does not impair repression of a *nanC::lacZ* reporter, whereas mutation of Spot 42 region I had a significant effect (Beisel et al. 2011 *Mol Cell*). This needs to be clarified.

We repeated the structural probing experiments on Spot 42 and on the *nanC* mRNA using unlabeled RNA concentrations close to the K_D of the Spot 42:*nanC* mRNA interaction. The new probing results demonstrate that the predicted pairing sites in Spot 42 and the *nanC* mRNA are more structured in the presence of the unlabeled RNA. The altered protection elsewhere in Spot 42 may be attributed to more extended base-pairing and/or a change in the secondary structure of Spot 42 upon hybridization with the *nanC* mRNA. We have added the following to p. 10 in the main text to address this:

> **"To assess whether Spot 42 can base pair with these targets through two regions, we performed in vitro structural probing with RNase T1, lead, and RNase V1 on Spot 42 complexed with the *nanC* mRNA. The altered cleavage patterns in the presence of unlabeled *nanC* mRNA supported base-pairing between regions I and III of Spot 42 and the *nanC* mRNA (Figure 3B). Altered cleavage also was observed outside of regions I and III, which may be attributed to more extended base-pairing and/or Spot 42 undergoing conformational changes upon pairing with the *nanC* mRNA."**

In our *Mol Cell* paper, there is evidence that mutation of region III reduces regulation of the *nanC* fusion (compare *nanC-I::lacZ* with pSpot42 to *nanC-I::laZ* with pSpot42-III in Figure 2B). The results support the contribution of both regions of Spot 42, where region I

contributes more than region III. We have added the following to p. 9 to address this:

> **"In many cases (e.g. *nanC*), one site predominantly contributed to regulation."**

5. How does Hfq impact the proposed Spot 42 interaction with multiple regions in targets?

We repeated the β-galactosidase assays in *hfq*-deletion strains on a natural fusion (*nanC*) and a synthetic fusion (*srlA+III*) that potentially interact with Spot 42 through multiple regions. In the original strain, both fusions were strongly repressed following Spot 42 overexpression. In the *hfq*-deletion strain, repression was compromised, reducing repression of the *nanC* fusion from 44-fold to 2.7-fold and the *srlA+III* fusion from 44-fold to 1.4-fold. Therefore, Hfq is important for target repression even when multiple targeting sites are present. The new data are reported in Supplementary Figure 3 and are noted in the main text on p. 9, 10.

6. Part of the data in Figure 4A is unclear. The parental *srlA::lacZ* fusion is regulated 2.8-fold in the presence of Spot 42 (as shown in Fig. 2B). Base-pairing occurs via region II of Spot 42, and mutation of this region impairs regulation (Beisel et al. 2011 *Mol Cell*). Now the authors extend the base-pairing with the 5'UTR of *srlA* (*srlA+III::lacZ*) and observe strongly increased regulation (27-fold repression), which is attributed to new base-pairing with region III of Spot 42. However, subsequent mutation of Spot 42 region III reverted repression not to 2.8-fold but to 6.3-fold. Moreover, the *srlA-II+III::lacZ* reporter (mutation of *srlA* opposite to part II of Spot 42) was still down-regulated by mutant pSpot42-III. In addition, in their previous article the authors showed that mutation of Spot 42 region III also impacted on the repression of the parental *srlA::lacZ* reporter (Beisel et al. 2011 *Mol Cell*). I know that none of these experiments and mutations is expected to give all-or-none results but this really needs to be explained much better.

One explanation for these observations is the remaining potential for base-pairing even after mutations were introduced into regions II and III of Spot 42. We have added the following on p. 10 to address this:

> **"The residual repression of *srlA-II+III::lacZ* by pSpot42-III may be attributed to a persisting potential for base-pairing even after mutations were introduced into sites II and III."**

From our *Mol Cell* paper, regulation of the *srlA::lacZ* fusion was slightly less for pSpot42-III than for pSpot42. We attributed this to reduced expression of this Spot 42 variant (*Mol Cell* paper, Figure 3B).

7. Figures 1, 2, 4, 5 and 6: the mutations introduced in the target mRNA fusions often impact on the basal β-galactosidase levels of the reporters. This indicates that translation activity and/or RNA stability greatly differs from the parental (wild-type) fusions. Since the authors use "fold-regulation" as a read-out for the base-pairing activity of Spot 42, they should show or argue why they ignore the possibility that altered translation rates (or ribosome occupancy) are responsible for the observed differences in regulation. Take the *srlA-entB::lacZ* reporter in Figure 6G, which has a basal β-galactosidase activity of ~8,000 MU and is regulated 4.2-fold by Spot 42. By contrast, the *entB+I::lacZ* reporter has a basal activity of only ~530 MU (more than a log difference) and shows only 1.7-fold regulation, despite longer base-pairing.

We agree that differences in mRNA stability or translational efficiency could have impacted the measured fold-change between different *lacZ* fusions. In most cases we only compared fold-changes for the same fusion (Figs. 1, 5, 6) or we only reported the presence (fold-change > 1.5) or absence (fold-change < 1.2) of regulation (Figs. 5, 6). In the few cases where we compared fold-changes between different fusions (Figs. 2,4), fold-changes varied more than basal levels. For instance, for the addition of a second pairing site to the *srlA* and *fucP* fusions, the fold-change increased by a factor of 9.6 for *srlA* and 2.8 for *fucP*, while basal levels increased by a factor of 0.94 for *srlA* and 1.2 for *fucP*.

8. Page 11, first para, and Figure 5: Does fusion of the 5' end of *srlA* to the predicted targets also impart association with Hfq in vivo? The in vitro experiments are suggestive, but looking into

enrichment of the "repaired" target fusion in coIP with Hfq in vivo/lysates, for example, using qRT-PCR, would be more convincing. This is a simple and straight-forward experiment. With regards to the in vitro work in Figure 5A, does a trimeric srlA-usg/Hfq/Spot 42 when the sRNA is added as well?

As suggested by this referee, we assessed Hfq binding in vivo using Hfq co-immunoprecipitation followed by primer extension analysis. Co-immunoprecipitation modestly enriched the *srlA* and *srlA-usg* fusion mRNAs in comparison to the *usg* fusion mRNA. No mRNA was co-immunoprecipitated in the *hfq*-deletion strain. These results support binding by Hfq in vivo following introduction of the 5' end of *srlA* onto the *usg* mRNA. The primer extension results have been substituted in place of the gel shift results in Figure 5C. The gel shift results now serve as supporting data and are included in Supplementary Figure 4. This is now addressed on p. 11 of the main text:

> **"The resulting *srlA-usg* mRNA was modestly enriched following co-immunoprecipitation of *E. coli* mRNAs bound to Hfq and showed increased binding to Hfq in vitro similar to that observed for *srlA* (Figure 5C, Supplementary Figure 4)."**

Because the in vivo Hfq co-immunoprecipitation assay results agreed with the in vitro gel shift assay results, we opted not to further investigate the interaction of Hfq with the RNAs in vitro.

9. Regarding guiding principles for base-pairing sRNAs, one would also mention a paper by Richter et al. 2010 Bioinformatics 26(1):1-5 in which a similar strategy discovered two novel targets of Yfr1 sRNA.

We thank the reviewer for pointing out this important reference. We now cite this article on p. 15 in the following sentence:

> **"Both *intaRNA* and *RNApredator* already account for the second criterion for target prediction, where use of *intaRNA* improved the identification of Ysr1 targets in *Prochlorococcus* MED4 (Richter *et al*, 2010)."**

10. Page 6, end of first para: ideally, one would like to see changes in abundance of the respective native target protein before concluding that some targets are false-positive. It could be that regulation is at the level or translation only, and/or the used reporters are just not suitable. Briefly mention these possibilities to caution.

We agree that generation of *lacZ* fusions could have compromised regulation of a real target gene. We have inserted the following sentence on p. 6:

> **"Negligible repression of the other eight reporters by Spot 42 suggests that these genes are not targets, although the possibility exists that generation of the *lacZ* fusions compromised regulation."**

11. Page 8, first para: again, regulation can be more than mRNA changes, which is what the cited papers assessed mainly.

We removed the Hao Y et al 2011 reference, which upon further inspection does not argue that more extended base-pairing leads to greater regulation. The remaining two references (Mitarai et al 2007, Mitarai et al 2009) are computational studies that do not differentiate between regulation at the levels of translation and mRNA stability.

12. Figure 1A: the structure model of Spot 42 needs labels to interpret the proposed base-pairing interactions.

We have numbered every ten nucleotides in Figure 1A to ease comparisons between the structural model and the predicted base-pairing interactions.

13. Figure 1: *puuE* is the third gene in an operon. Does Spot 42 regulate the upstream *puuC* and

puuB genes as well?

We employed *TargetRNA* to search for potential base-pairing interactions between the three unstructured regions of Spot 42 and the upstream *puuC* and *puuB* genes. No substantial interactions were predicted by *TargetRNA*, suggesting that Spot 42 targets only one gene in the *puuCBE* operon.

14. Table II: the authors argue on page 7, 2nd para, that 10/15 predicted targets were repressed. However, Table II includes only nine (*nanT*, *nanC*, *paaK*, *atoD*, *ascF*, *galK*, *srlA*, *fucP* and *glpF*).

The tenth gene showing regulation is *caiA*. Although the fold-regulation for *caiA* (1.6) was lower than that for the other targets, this value was consistently above the maximum fold-change observed with the empty vector pBRplac (1.2).

15. Figure 2B: for the *srlA_L::lacZ* extended base-pairing into the structured Spot 42 sequence improved base-pairing from 2.8-fold to 5.5-fold while *srlA_R::lacZ* (extending into the unstructured regions of Spot 42) improved base-pairing to 12.5-fold. Is this really due to different stabilities of complexes or to different structural constraints (the Spot 42-*srlA_L* interaction will require three additional bulged residues)?

The three-residue bulge is the start codon, which could not be altered in *srlA_L* without jeopardizing the translational efficiency of the fusion. Since regulation of *srlA_L* could be explained in part by the interrupted pairing, we have altered part of p. 8 to read the following:

> "In contrast, extending base-pairing into the structured region of *gltA* (*gltA_R*) and fucP (*fucP_L*) did not improve regulation. Extending base-pairing into the structured region of *srlA* (*srlA_L*) improved regulation less than what was observed when base-pairing was extended through the remainder of the unstructured region (*srlA_L*), although interpretation of this result is complicated by the necessity of having the start codon interrupt the extended pairing."

16. Fig. 6G: the increase in β-galactosidase activity of the *entB+I::lacZ* reporter when transformed with the Spot 42 plasmid (without IPTG) should be explained, specifically because this is crucial for the interpretation of the experiment.

We repeated the β-galactosidase assay on the *entB+I* fusion harboring pSpot42. Although basal levels varied between assays (738 MU versus 327 MU), the fold-regulation was very similar (1.7 versus 1.8). We have introduced the new data in Figure 6G. A caveat about some variability in the basal levels has been added to the Figure 6 legend.

Referee #2

1. Since the experiments in Figs 5 and 6 involve major mutational changes, which could have effects other than the ones desired, and since Hfq binding is here assumed to be a key factor in reactivating silent targets, these experiments should be repeated in an Hfq-deficient strain. This straightforward experiment should also be done for the extended-target fusions in Figure 2, in particular since the Gottesman lab has argued that longer interactions may overcome an Hfq requirement.

As suggested by this referee, we repeated the β-galactosidase assays on *hfq*-deletion strains harboring fusions selected from Figure 2 (*srlA_R*), Figure 5 (*srlA-usg*), and Figure 6 (*srlA-moeA1,2* and *srlA-entB*) as well as the *nanC* and *srlA* fusions. Deletion of *hfq* compromised repression of all the fusions by Spot 42. The new data are reported in Supplementary Figure 3 and are noted on p. 11, 12, and 14 in the main text.

2. Two of the three not regulated cases (*usg*, *moeA*) are straightforward. The *entB* case is complicated, and even though the interpretations given may turn out to be correct, there are loose ends. Since competition between Hfq and Spot 42 for the downstream Hfq binding

site/target site is assumed, this could easily be tested in vitro. For instance, by analyzing gelshifts of an mRNA fragment lacking the upstream binding site with Spot 42 and/or Hfq. And again, an in vivo test in an Hfq mutant strain might help. It might also be worth a few more lines to discuss whether such a model can account for the in vivo effects. For Hfq blocking access to a target site, its occupancy must be very high (most of the mRNAs at any given time bound by Hfq at this site).

As suggested by this referee, we repeated the β-galactosidase assay on the *srlA-entB* fusion in an *hfq*-deletion strain. The results now are included in Supplementary Figure 3 and are noted on p. 14 in the main text:

> **"In addition, regulation of the *srlA-entB* fusion by Spot 42 was disrupted when *hfq* was deleted (Supplementary Figure 3), suggesting that Hfq is required for Spot 42 to associate with the *entB* mRNA in vivo."**

We opted not to perform the suggested in vitro gel shift assay with Hfq, Spot 42, and the *entB* mRNA because of anticipated complications interpreting the results. While our competition model might imply that Hfq, Spot 42, and the *entB* mRNA cannot form a ternary complex in vitro, Hfq can bind Spot 42 and the *entB* mRNA independently. This would allow formation of a ternary complex even if Spot 42 and the *entB* mRNA are unable to hybridize.

3. Fig. 3: Though the reporter gene experiments are convincing, the biochemical analysis of the interaction site is not. Except for the lead cleavages at 2 μM *nanC*, I find it difficult to see any convincing pattern. In particular, it is unclear why RNAse III was used as a probe for dsRNA regions. V1 has this specificity whereas RNAse III usually cleaves 20 or so base-pair stretches of RNA, far longer than assumed here (Fig 3A). Incidentally, the choice of concentration ranges in this experiment, as stated, comes from the binding curves in Suppl. F.3. There, the mid-point (50% labeled RNA bound) differs in the reciprocal experiments by maybe 4-fold. This is odd, to say the least, unless a binding mode different from 1:1 is assumed.

As suggested by this referee, we repeated the entire structural probing experiment for Spot 42 and for the *nanC* mRNA using RNase V1 in place of RNase III. The most significant changes observed in these probing experiments were in the two predicted regions of pairing.

We only performed the gel shift experiments to estimate suitable concentrations of unlabeled RNAs for the structural probing experiments. The differences in K_D's between the gel shift assays likely reflect experimental error as well as differences in the folding properties of the RNAs.

4. For many of the fusion constructs, adding for example the srlA leader region changes lacZ activities substantially (e.g. moeA down 5x, entB up 8x). Is this an indication of changed mRNA levels, or translational activity?

As suggested by this referee, we measured the relative mRNA levels of the *moeA* and *entB* fusions with or without the 5' end of *srlA* by primer extension analysis. For both fusions, the addition of the 5' end of *srlA* resulted in similar changes in β-galactosidase activities and in mRNA levels. This suggests that differences in β-galactosidase activities reflect differences in mRNA levels, although alterations in translational activity cannot be ruled out. The new data are included in Supplementary Figure 7 and are noted in the legend to Figure 6:

> **"Different fusions showed differing levels of basal expression, which for *moeA* and *srlA-moeA* (D) as well as *entB* and *srlA-entB* (G) are reflected in differing mRNA levels (Supplementary Figure 7)."**

Referee #3

1. Many targets predicted by "TargetRNA" using the full-length Spot 42 seem to base pair with the unstructured regions of Spot 42 (Table 1). Are these unstructured regions different from the regions I, II, and III or do they overlap each other?

The unstructured regions of Spot 42 designated in Table I, Table II, and Figure 1A are the same. This is now clarified in the legend of Table I:

> **"Nucleotides highlighted in gray designate the unstructured regions of Spot 42 (Figure 1A)."**

2. The mRNA levels of two targets regulated as *lacZ* fusion by Spot 42 are significantly elevated in the Δ*spf* strain lacking Spot 42 as expected while those of others are not (Supplementary Fig. 1). What are the possible reasons for this?

We see two potential reasons for the lack of upregulation in the targets validated as *lacZ* fusions: the target mRNAs are regulated only at the level of translation, or endogenous levels of Spot 42 are insufficient to show measurable regulation of these targets.

We have added the following to p. 7:

> **"Among the genes tested, two of the five regulated as *lacZ* fusions (*glpF*, *paaK*) were significantly upregulated in the Δ*spf* strain. The other three genes may be regulated at the level of translation or are not measurably regulated by Spot 42 under the conditions tested."**

2nd Editorial Decision 06 February 2012

Thank you for sending us your revised manuscript.

Referees 1 and 3 have now seen the manuscript again, and you will be pleased to learn that in their view you have addressed all criticisms in a satisfactory manner. The paper will now be publishable in The EMBO Journal.

Prior to formal acceptance, there are two editorial issues that need further attention:

You may wish to follow the suggestion of referee 1 to increase the font size for the nucleotides shown in figure 6A.

We now encourage the publication of source data, particularly for electrophoretic gels and blots, with the aim of making primary data more accessible and transparent to the reader. Would you be willing to provide files comprising the original, uncropped and unprocessed scans of all gels used in the figures? We would need 1 file per figure (which can be a composite of source data from several panels) in jpg, gif or PDF format, uploaded as "Source data files". The gels should be labelled with the appropriate figure/panel number, and should have molecular weight markers; further annotation would clearly be useful but is not essential. These files will be published online with the article as a supplementary "Source Data". Please let me know if you have any questions about this policy.

Please follow the link below to upload the files.

Thank you very much again for considering our journal for publication of your work.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1

The authors have satisfactorily addressed my previous criticism.

One minor issue for the production process: Please increase font size in Fig 6A.

Referee #3
The manuscript has been substantially improved by addressing most of the criticisms/concerns raised by three referees including me. I believe the revised manuscript could be published in EMBO J.

2nd Revision - authors' response 07 February 2012

We thank the referees for reviewing the revised version of our manuscript.

Referee #1

1. The authors have satisfactorily addressed my previous criticism. One minor issue for the production process: Please increase font size in Fig 6A.

As suggested by the referee, we increased the font size of the text in Fig. 6A.

Referee #3

The manuscript has been substantially improved by addressing most of the criticisms/concerns raised by three referees including me. I believe the revised manuscript could be published in EMBO J.

